# 人类语音起源和演化研究综述

孔江平 吴西榆

**摘要** 本文详细地梳理了人类语音起源和演化的研究成果，为对语音起源和演化感兴趣的学者和研究生提供了大量有关的研究文献，并对前人在这个领域里的研究进行了初步的评述。

**关键词** 语音的起源 语音的演化 喉头下降理论 灵长类的发音 古人类化石

# A review on the origin and evolution of human speech

**Kong Jiangping Wu Xiyu**

**Abstract** This paper reviews the research results of the origin and evolution of human speech in detail, provides a large number of relevant literature for scholars and graduate students who are interested in the origin and evolution of human speech, and gives a preliminary review of previous studies in this field.

**Keywords:** the origin of speech; the evolution of speech; the theory of laryngeal descent; the vocalization of primates; fossil hominins



## 1.引言

语言在人类的演化过程中起着重要的作用，如何找到证据来揭示语言的起源是古人类学家、哺乳动物学家、语音学家、语言学家和言语科学家一直追求的目标。语言起源的研究是一个多学科交叉的科学，它整合了古人类学、解剖学、语音学、言语科学、认知科学等相关科学。在语言演化的研究中，言语发音的研究最具有可行性，因为可以采用实证的科学方法。从目前的研究看，占主导地位的是 Lieberman 的喉头下降理论（LDT）。Lieberman 首先从生物学的角度研究了言语的演化，后来他又从：人类言语、现有非人类灵长类的语言能力、人类言语演化的化石记录和喉上声道（SVT）、言语的神经机制等方面讨论了言语演化的过程，最终提出了喉头下降的理论（[1]Lieberman,1984; [2]Lieberman et al,2015）。Lieberman 的喉头下降理论在学界几乎主导了半个世纪。然而，在近十年，由于新的研究证据的积累和观点的提出，喉头下降理论受到了挑战（[3]Fitch，2000）。这些新的证据和争论主要集中在：1）哺乳动物和非人类灵长类喉头的下降是否能产生人类的语音（[4]Fitch et al, 2001）？2）现在存活的非人类灵长类动物的声道是否是统一形状？在喉头不下降的情况下，它们是否可以发出具有共振峰对立模式的声音？（[5]Boë et al，2017)。3）在研究方法上，利用声道模型模拟元音的演化和涌现是否有效？4）怎样界定元音的涌现和涌现的时间？Boë 等学者对这些争论、新的观点和研究方法做了很好的评论（[6]Boë et al，2019）。下面我们将对人类语音的起源和演化研究做一些梳理和综述。

## 2．喉头下降理论

上个世纪 50 年代，Lieberman 就开始利用新的技术和概念研究非人类灵长类的发音，是这一研究领域的开创者之一。从上个世纪 70 年代开始，Lieberman 等利用频谱分析的方法，分析了猕猴、黑猩猩和大猩猩的叫声，并根据声学分析的结果研

究了声道解剖和声学特征之间的关系。认为它们的发音主要是类似央元音的声音，暗示了非人类灵长类在叫喊时没有能力改变喉上声道的形状（[7]Lieberman，1968）。由于猕猴和猿类咽腔很小，它们不能发出人类的语音，因为这是本质的生理制约，所以无法产生连续变化的语言所需的声道面积（[8]Lieberman et al，1969）。Lieberman 还研究了出生婴儿的哭声，发现人类婴儿同其他非人类灵长动物一样，除了喉头整体运动，在说话时不会使用喉上声道。由于受到发音器官的限制，他们也缺少咽腔面积的变化（[9]Lieberman et al,1971）。Lieberman 和 Crelin 还利用化石模型来推断尼安德特人声道的解剖结构和声道形状，用于估算尼安德特人的发音潜力，最终将这种发音潜力与人类的语音进行了对比研究（[10]Lieberman et al, 1971）。Lieberman 的研究表明,声道的功能不只是为了呼吸和吞咽，因为人类演化的骨骼证据显示出了一系列的变化，这些变化可能是灵长类动物的声道在某种程度上为了能够说话而形成的。因此，应被视为人类的语音输出机制的一部分，即人类特有的语言能力。

另一项对声道演化的重要研究是从古人类化石的角度。黑猩猩和现代人类的分化可以追溯到大约 6 百万年前（[11]Vallender Eric, et al. 2008.），由于获取化石资料十分困难，因此，从古人类的化石研究声道的演化受到了制约。从目前世界主要的古人类化石看，同时具有头颅(cranium, skull)、上颌(upper mandible)、上齿(upper teeth)、下颌(lower mandible)、下齿(lower teeth)的化石只有十几种，分别是：Pataud 1、Combe-Capelle、Dolni Vestonice XV and XVI、Barma Grande 5、a Ferrassie 1、Le Moustier、Pech De L'aze、Roc de Marsal、Fish Hoek、Nariokotome、Amud、Jebei Qafzeh 9、Jebei Qafzeh11、Shanidar 1、Skhul V、Tabun I、Teshik-Tash、Zhoukoudian Upper Cave: Cast of PA.101。但古人类完整头颅的化石相对较多，目前大约几十种（[12]Jeffrey et al, 2002; [13]Jeffrey et al,, 2003）。由于口腔中的肌肉和喉头的软骨不会形成化石，这为重构声道的研究带来了困难。

在利用化石对人类声道演化的研究中，研究人员首先测量比较了现代人类婴儿和成人喉头下降的发展过程（[14] Laitman et al，1976），同时测量研究了灵长类动物头颅化石和人属头颅化石的颅基线上的牙槽中点（A, prosthion）、口后点（B, staphylion）、蝶枕点（C, hormion）、蝶骨点（D, sphenobasion）和颅底点（E, endobasion）。研究发现：咽上收缩肌向前和向下几乎垂直于枕骨平面的定义线条 DE，并比较研究了人属灵长类和非人属灵长类颅骨上六点的距离，讨论了这些距离和喉头下降的关系，（[15] Laitman et al，1978；[16] Laitman et al，1979; [17] Laitman et al，1982）。根据以上的这些对化石的研究，Budil 发展出了一种计算这些参数的方法，提出了 K 参数，并对南猿、尼安德特人和智人的头颅参数进行了 K 参数计算，并在此基础上对人属的声道进行了功能重构（[18] Budil，1994）。

Liberman 等对黑猩猩、现代人类婴儿及成人发音器官生理形态和发育过程的研究以及 Laitman 等对人属化石的研究，都对 Lieberman 的喉头下降理论（LDT）提供了证据。在这一理论框架下，人们对黑猩猩到现代人类的声道演化进行了大量研究，并对古人类声道进行了重构，认为喉头下降是人类声道演化和产生语音最重要的演化途径（[1] Lieberman, 1984; [2] Lieberman et al, 2015）。基于声学和解剖观察的坚实基础,Lieberman 的结论被认为是事实,是这一研究领域的试金石,最终被称为喉头下降理论(LDT)。

3. 非人类灵长类和哺乳动物喉头的下降

对哺乳动物喉头的研究发现，有些哺乳动物的喉头也会下降，如鹿（[4] Fitch et al，2001）、蒙古羚羊（[19] Frey et al，

2003；）、狮子、美洲虎、老虎、猎豹和家猫（[20] Weissengruber et al，2002）、黑猩猩（[21] Nishimura et al，2003; [22] Nishimura, 2005）、灵长类（[23] Nishimura，2003）。并对非人类灵长类发音的连续性和不连续性进行了研究（[24] Berthommier,et al,2018）。Fitch 发现并记录了哺乳动物山羊、狗、猪和绢毛猴的低喉头、舌骨和舌头更广泛的动态和临时的过程，认为喉头下降不是人类独有的，并暗示了喉头下降是两条演化的不同路径（[3] Fitch，2000）。他们认为喉头下降是为了延长声道，通过降低声道的共鸣频率让人感知到被夸大的体型，因此，体型夸大提供了一个非语言学的替代假设([25] Fitch, 1997)。他认为像所有哺乳动物一样，最早的人科动物在发音时重新配置发音解剖结构就可以获得足以产生任何清晰易懂音素的声道结构。但可以看出非人类动物喉头下降所形成的声道形态与人类的声道形态还是有很大的差异。

4. 非人类灵长类的发音

通常认为灵长类动物中只有人类的发音器官具有语言发音能力，并默认包括非人类灵长类动物在内的哺乳动物的声道在发声时类似于一个均匀统一的（uniform）或张开的管道([7] Lieberman,1968; [8] Lieberman et al, 1969；[26] Shipley et al, 1991)。但近期的研究表明非人类灵长类也有可能发出不同的元音。对南非狒狒的发音研究表明，狒狒的声道不是均匀的 ([27] Owren et a, 1997)。Riede 等人首先测试了 Diana monkey 的报警叫声与管道特性的一致性，他们的研究结果表明 Diana monkey 对鹰和豹子（baoleopard）的报警叫声并不像 schwa，而是类似于 10 至 12 岁儿童/a/的发音([28] Riede et al,2005)。研究结果与 VTL 相似的儿童发出的人类元音相关（[29] Lee,1999）。Fitch 等人研究了猕猴有可能发出的元音，认为猕猴的声道可以很容易地产生足够的语音范围来支持口语，是可以发出人类的语音的（[30] Fitch et al，2016）。Boë 和他的同事们对狒狒(Papio Papio)的研究表明，狒狒是可以自然产生类似人类共振峰结构的元音（[5] Boë et al，2017)。Pisanski 和他的同事认为这可能是早期声音控制能力的活的遗留物，这种能力导致了人类语言的清晰表达([31]Pisanski et al,2016)。Fitch 将猕猴不能发音归于猕猴还没有演化出能够产生语言的神经系统，认为人类语言能力的进化需要的是神经系统的变化，而不是声音结构的改变（[30] Fitch et al，2016）。基于对非人类灵长类声道解剖和发音的研究，有一些学者认同“神经假说”，认为非人类灵长类的发音器官在发元音上没有问题，问题是它们缺乏相应的脑神经机制（[32]Boer，2010; [33]Boer et al, 2010; [34] Kelly, 2016）。但 Lieberman 认为，Fitch 等人研究中的言语应该是猕猴的言语发音，并不能将这个言语看成是完全意义上的人类言语（[35]Lieberman，2017）。

5. 声道模型研究

有些学者认为生物学家在试图将语言研究的原则和方法应用于动物交流时遇到了困难（[30] Fitch et al,2016），根据不同声道模型和言语产生研究的新成果，人们对声道和言语产生的关系有了更多的认识。在 Lindblom 和 Harshman 等人研究的基础上([36]Lindblom et al, 1971；[37]Harshman et al, 1977), Maeda 的模型提高了发音声道使用主成分分析建模的矢状变量，Maeda 用 7 个控制参数可以解释 70%-90%的元音变化（[38]Maeda，1990）。利用这一模型和肌电信号，人们研究了声道模型控制参数和舌头肌肉的活动（[39]Maeda et al,1994；[40] Honda, 1996）以及生理机制（[41]Buchaillard et al,2009）。Goldstein 将大量的解剖学数据整合到一个清晰的声道生长模型中([42]Goldstein,1980)，模型采用了水平声道长度、垂直声道长度和声道中线长度三个参数。Goldstein 在 Mermelstein 成人模型的基础上([43]Mermelstein, 1973)，利用她的数据集

建立了声道矢状面生长模型。给出了声道的面积函数和以此产生的共振峰值模拟研究新生儿的发音，认为新生儿能发出元音/i/，/a/,/ u /。Boë 的实验室开发了可变线性关节模型(VLAM)来模拟声道从出生到成年的生长过程([44]Boë et al，1997)，声道长度的生长模式引入了前口腔和后咽腔两个定标因子，中间区域则是在这两个定标因子之间进行插值。研究表明，即使婴儿和儿童的咽腔长度不到成年体长度的一半，他们也能在喉头下降之前就开始形成元音对立。Badin 的研究证明嘴唇的动作可以补偿喉头下降的作用，即使在高喉的情况下，也能产生正常的 F1-F2 空间([45]Badin et al,2014)。Boë 等认为无论咽腔多长，从出生到成年，咽管的大小和形状都是相当稳定的([46]Boë et al,2013)。但 Boer 认为这些模型过于简化([32]Boer，2010)。从声道模型的研究可以看出，不断改进声道模型可以推进声道演化的模拟研究。

6. 元音涌现时间的推论

元音的涌现主要包括两个方面：一是元音涌现的方式和过程；二是元音涌现的时间。目前对于元音涌现的条件和过程，主要是利用古人类的化石信息构建发音生理形态和模型来讨论元音的涌现。另一个方面是通过研究人类和非人类灵长类的发音来讨论元音的涌现。这些研究本文在前面几节已经做了充分的综述。在元音涌现的时间上，可以分为一下几种观点：1）Lieberman，Laitman，Crelin 等学者进行了大量的化石研究，认为尼安德特人不具有现代人类的发音能力，从发音形态上只具有较低的言语能力。也有学者认为现代人了和尼安德特人不属于同一支，因此，不可以类比。Lieberman 等学者认为言语大约在直立人后期出现（0.1 mya）（[1]Lieberman,1984; [2]Lieberman et al,2015; [15]Laitman et al,1978; [47] Laitman,1985；[48] Crelin, 1989）。2)Budil 根据 Laitman 等对古人类头颅基的化石的测量数据，发展了一种 K 系数。并在此基础上进行了古人类发音形态的重构研究，推测古人类的发音应该形成于上更新世（the Upper Pleistocene, 0.126-0.0117 mya）（[18]Budil, 1994）。3）Arensburg 对出土的一个完整的舌骨化石进行了研究，认为尼安德特人的舌骨和发音形态与智人已经没有差别。他推测言语涌现的时间在旧石器时期中期（Middle Palaeolithic，大约 0.2 to 0.045 mya）。发音器官的演化应该更早，在旧石器中期之前已经开始演化和形成([49]Arensburg et al, 1989）。4) Beo 等学者对当前言语演化和起源的研究进行了梳理，认为现有的证据压倒性地驳斥了长期存在的喉部下降理论，将"语言的起源"从大约 0.2 mya 推到大约 20 mya（[6]Boë et al, 2019）。

7. 结语

从以上的研究可以看出，无论是喉头下降理论还是对这种理论的质疑和新的证据，关键问题主要集中：1）喉头下降是否就一定能产生语言所需的三个基本元音；2）喉头不下降是否也能产生三个语言的基本元音；3）元音共振峰的对立模式怎样来判定；4）元音涌现时间的推论。因此，喉头下降理论和新的研究对这一理论的质疑和争论构成了现在人类言语声道演化研究的现状。

## 参考文献

[1] Lieberman, P., 1984, The Biology and Evolution of Language, Cambridge, Massachusetts: Harvard University Press.

[2] Lieberman and McCarthy, 2015, The Evolution of Speech and Language, Handbook of Paleoanthropology, W. Henke, I. Tattersall (eds.), Springer-Verlag Berlin Heidelberg。

[3] W. T. Fitch, 2000，The phonetic potential of nonhuman vocal tracts: Comparative cineradiographic observations of vocalizing animals. Phonetica 57, 205–218.

[4] W. T. Fitch, D. Reby, 2001，The descended larynx is not uniquely human. Proc. R. Soc. Lond. B Biol. Sci. 268, 1669–1675.

[5] Boë L.-J., F. Berthommier, T. Legou, G. Captier, C. Kemp, T. R. Sawallis, Y. Becker,A. Rey, J. Fagot, 2017, Evidence of a vocalic proto-system in the baboon (Papio papio) suggests pre-hominin speech precursors. PLOS ONE 12, e0169321.

[6] Boë L-J, Thomas R. Sawallis, Joël Fagot, Pierre Badin, Guillaume Bartier, Guillaume Captier, Lucie Ménard, Jean-Louis Heim, Jean-Luc Schwartz, 2019, Which way to the dawn of speech?: Reanalyzing half a century of debates and data in light of speech science, Science Advances, American Association for the Advancement of Science, 1200 New Science York Avenue NW, Washington, DC 20005.

[7] P. Lieberman, Primate vocalizations and human linguistic ability, 1968. J. Acoust. Soc. Am. 44, 1574–1584.

[8] P. Lieberman, D. H. Klatt, W. H. Wilson, 1969, Vocal tract limitations on the vowel repertoires of rhesus monkey and other nonhuman primates. Science 164, 1185–1187.

[9] P. Lieberman, K. S. Harris, P. Wolff, L. H. Russel, 1971, Newborn infant cry and nonhuman primate vocalization. J. Speech Hear. Res. 14, 718–727 .

[10] P. Lieberman, E. S. Crelin, 1971, On the speech of Neanderthal man. Linguist. Inq. 2, 203–222.

[11] Eric J. Vallender， Nitzan Mekel-Bobrov， Bruce T. Lahn, 2008, Genetic basis of human brain evolution, Trends in Neuroscience.

[12] Jeffrey H. Schwartz, Ian Tattersall, 2002, THE HUMAN FOSSIL RECORD, Volume One, Craniodental Morphology of Genus Homo, (Africa and Asia). WILEY-LISS, John Wiley& Sonns,Inc..

[13] Jeffrey H. Schwartz, Ian Tattersall, 2003, THE HUMAN FOSSIL RECORD, Volume Two, Craniodental Morphology of Genus Homo, (Africa and Asia). WILEY-LISS, John Wiley& Sonns,Inc..

[14] Laitman JT, Crelin ES (1976) Postnatal development of the basicranium and vocal tract region in man. In: Bosma JT (ed) Development of the basicranium. National Institutes of Health, Bethesda

[15] Laitman JT, Heimbuch RC, Crelin ES (1978) Developmental change in a basicranial line and its relationship to the upper respiratory system in primates. Am J Anat 152:467–482

[16] Laitman JT, Heimbuch RC, Crelin ES (1979) The basicranium of fossil hominids as an indicator of their upper respiratory systems

[17] Laitman JT, Heimbuch RC (1982) The basicranium of Plio-Pleistocene hominids as an indicator of their upper respiratory systems. Am J Phys Anthropol 59:323–343

[18] Ivo Budil, 1994, Functional Reconstruction of the Supralaryngeal Vocal Tract of Fossil Human, HUMAN EVOLUTION, Vol. 9- N. 1 (35-52).

[19] R. Frey, T. Riede, 2003，Sexual dimorphism of the larynx of the Mongolian gazelle (Procapra gutturosa Pallas, 1777) (Mammalia, Artiodactyla, Bovidae). Zool. Anz. J. Comp. Zool.242, 33–62 (2003).

[20] G. E. Weissengruber, G. Forstenpointner, G. Peters, A. Kübber-Heiss, W. T. Fitch, 2002， Hyoid apparatus and pharynx in the lion (Panthera leo), jaguar (Panthera onca), tiger(Panthera tigris), cheetah (Acinonyx jubatus) and domestic cat (Felis silvestris f. catus).J. Anat. 201, 195–209.

[21] Takeshi Nishimura, Akichika Mikami, Juri Suzuki, and Tetsuro Matsuzawa (2003) Descent of the larynx in chimpanzee infants, PANS, June 10, vol. 100, no. 12: 6930-6933.

[22] Takeshi Nishimura, 2005, Developmental Changes in the Shape of the Supralaryngeal Vocal Tract in Chimpanzees, AMERICAN JOURNAL OF PHYSICAL ANTHROPOLOGY 126:193–204

[23] F. Berthommier, L.-J. Boë, A. Meguerditchian, T. R. Sawallis, G. Captier, 2018, In Origins of Human Language: Continuities and Discontinuities with Nonhuman Primates, L.-J. Boë, J. Fagot, P. Perrier, J.-L. Schwartz, Eds. (Peter Lang, 2018), pp. 101–135.

[24] W. T. Fitch, Vocal tract length and formant frequency dispersion correlate with body sizein rhesus macaques. J. Acoust. Soc. Am. 102, 1213–1222 (1997).

[25] Carl Shipley, Edward C. Carterette, J S Buchwald, 1991, The effects of articulation on the acoustical structure of feline vocalizations, Article in The Journal of the Acoustical Society of America 89(2):902-9

[26] M. J. Owren, R. M. Seyfarth, D. L. Cheney,1997, The acoustic features of vowel-like grunt calls in chacma baboons (Papio cyncephalus ursinus): Implications for production processes and functions. J. Acoust. Soc. Am. 101, 2951–2963.

[27] T. Riede, E. Bronson, H. Hatzikirou, K. Zuberbühler, 2005, Vocal production mechanisms in a non-human primate: Morphological data and a model. J. Hum. Evol. 48, 85–96.

[28] Lee, A. Potamianos, S. Narayanan, 1999，Acoustics of children’s speech: Developmental changes of temporal and spectral parameters. J. Acoust. Soc. Am. 1051455–1468.

[29] W. T. Fitch, B. Boer, N. Mathur, A. A. Ghazanfar, 2016, Monkey vocal tract are speech-ready, Sci. Adv. 2, e1600723.

[30] K. Pisanski, V. Cartei, C. McGettigan, J. Raine, D. Reby, 2016，Voice modulation: A window into the origins of human vocal control? Trends Cogn. Sci. 20, 304–318.

[31] B. de Boer, 2010，Modelling vocal anatomy’s significant effect on speech. J. Evol. Psychol. 8,351–366.

[32] B. de Boer, W. T. Fitch, 2010, Computer models of vocal tract evolution: An overview and critique. Adapt. Behav. 18, 36–47.

[33] B. de Boer, 2010, Investigating the acoustic effect of the descended larynx with articulatory models. J. Phon. 38, 679–686.

[34] M. Kelly, 2016, Monkey Speak: Macaques Have the Anatomy, Not the Brain, for Human Speech (Princeton University News, 2016); www.princeton.edu/ news/2016/12/09/ monkeyspeak-macaques-have-anatomy-not-brain-human-speech).

[35] Philip Lieberman，2017, Comment on monkey vocal tract are speech-ready, Sci. Adv. 3, e1700442 (2017).

[36] B. Lindblom, J. Sundberg, 1971, Acoustical consequences of lip, tongue, jaw, and larynx movement. J. Acoust. Soc. Am. 50, 1166–1179.

[37] R. Harshman, P. Ladefoged, L. Goldstein, 1977, Factor analysis of tongue shapes. J. Acoust. Soc. Am. 62, 693–707.

[38] S. Maeda, 1990, In Speech Production & Speech Modelling, W. J. Hardcastle, A. Marchal, Eds. (Kluwer, Dordrecht, 1990), pp. 131–149.

[39] S. Maeda, K. Honda, 1994， From EMG to formant patterns of vowels: The implication of vowel spaces. Phonetica 51, 17–29.

[40] K. Honda, 1996, Organization of tongue articulation for vowels. J. Phon. 24, 39–52.

[41] S. Buchaillard, P. Perrier, Y. Payan, 2009, A biomechanical model of cardinal vowel production: Muscle activations and the impact of gravity on tongue positioning. J. Acoust. Soc. Am.126, 2033–2051.

[42] Ursula Gisela Goldstein, 1980, “An articulatory model for the vocal tracts of growing children,” D. Sc. thesis, Massachusetts Institute of Technology.

[43] P. Mermelstein, 1973，Articulatory model for the study of speech production. J. Acoust. Soc. Am. 53, 1070–1082

[44] L.-J. Boë, S. Maeda, 1997，in Journées d’Études Linguistique (Université de Nantes,pp. 98–105.

[45] P. Badin, L.-J. Boë, T. R. Sawallis, J.-L. Schwartz, 2014, Keep the lips to free the larynx: Comments on de Boer’s articulatory model (2010). J. Phon. 46, 161–167.

[46] L.-J. Boë, P. Badin, L. Ménard, G. Captier, B. Davis, P. MacNeilage, T. R. Sawallis,J.-L. Schwartz, 2013, Anatomy and control of the developing human vocal tract: A response to Lieberman. J. Phon. 41, 379–392.

[47] Laitman. J.T. in Hominid Evolution, Past Present and Future (ed. Tobias, P.V.)181-186 (Liss, New York, 1985).

[48] E. S. Crelin， 1989, The Human Vocal Tract: Anatomy, Function, Development and Evolution，Vantage New York.

[49] B. Arensburg, A.M. Tillier, B. Vandermeersch, H. Duday, L.A.Schepartz & Y. Rak, 1989, A Middle Palaeolithic human hyoid bone, Nature, Vl. 338, 27 APRIL.

**孔江平** 北京大学中文系，博士，教授，主要研究领域为语音学、言语产生和语音的起源。
Email: jpkong@pku.edu.cn

**吴西榆** 北京大学中文系，博士，助理教授，副研究员，主要研究领域为言语产生、语音生理模型和言语障碍。
Email: xiyuwu@pku.edu.cn

**以下为其他有关文献**

Duchin, L. E. (1990). The evolution of articulate speech: Comparative anatomy of the oral cavity in Pan and Homo. Journal of Human Evolution, 19, 687–697. (see attached)

Takemoto H. (2008) Morphological analyses and 3D modeling of the tongue musculature of the Chimpanzee (Pan troglodytes). Am J Primatol 70:966–975

S. Adachi, M. Yamada, 1999, An acoustical study of sound production in biphonic singing, Xo¨o¨mij, J. Acoust. Soc. Am. 105 (5).

Dennis H. Klatt， 1980， Software for a cascade/parallel formant synthesizer， J. Acoust. Soc. Am. 6/(3). Mar. 1980.

Alvin M. Liberman, Katherine Safford Harris, Howard S. Hoffman, and Belver C. Griffith, 1957, The discriminnation of speech sounds within and across phoneme boundaries, Journal of Experimental Psychology, Vol. 54, No. 5, 1957.

Liberman, A. M. 1957. Some Results of Research on Speech Perception. The Journal of the Acoustical Society of America, 29(1), 117-123.

B. de Boer, 2010, Investigating the acoustic effect of the descended larynx with articulatory models. J. Phon. 38, 679–686.

P. Lieberman, The evolution of human speech: Its anatomical and neural bases. Curr. Anthropol. 48, 39-66 (2007).

P. Lieberman, The evolution of language and thought. J. Anthropol. Sci. 94, 127-146, (2016).

W. T. Fitch, R. A. Suthers, 2016, In Vertebrate Sound Production and Acoustic Communication, R. A. Suthers, W. T. Fitch, R. R. Fay, A. N. Popper, Eds. (Springer,), pp. 1–18.

B. Lindblom, J. Sundberg, 1969， A quantitative model of vowel production and the distinctive features of Swedish vowels. Speech Transm. Lab. Q. Prog. Status Rep. 10, 14–32.

Crelin, E.S., 1987, The Human Vocal Tract, Vantage New York 1987.

Ian Tattersall, 2008, An Evolutionary Framework for the Acquisition of Symbolic Cognition by Homo sapiens, Comparative Cognition & Behavior Reviews. Vol. 3, pp 99-114.

T. Nishimura, 2003，Comparative morphology of the hyo-laryngeal complex in anthropoids: Two steps in the evolution of the descent of the larynx. Primates 44, 41–49 (2003).

D. Rendall, S. Kollias, C. Ney, P. Lloyd, 2005, Pitch (F0) and formant profiles of human vowels and vowel-like baboon grunts: The role of vocalizer body size and voice-acoustic allometry. J. Acoust. Soc. Am. 117, 944–955.

H. K. Vorperian, S. Wang, E. M. Schimek, R. B. Durtschi, R. D. Kent, L. R. Gentry,M. K. Chung, 2011, Developmental sexual dimorphism of the oral and pharyngeal portions of the vocal tract: An imaging study. J. Speech Lang. Hear. Res. 54, 995–1010.

H. K. Vorperian, S. Wang, M. K. Chung, E. M. Schimek, R. B. Durtschi, R. D. Kent, A. J. Ziegert, L. R. Gentry, 2009, Anatomic development of the oral and pharyngeal portions of the vocal tract: An imaging study. J. Acoust. Soc. Am. 125, 1666–1678 (2009).

L. Ménard, B. L. Davis, L.-J. Boë, J.-P. Roy, 2009, Producing American English vowels during vocal tract growth: A perceptual categorization study of synthesized vowels. J. Speech Lang. Hear. Res. 52, 1268–1285 (2009).

L. Ménard, J.-L. Schwartz, L.-J. Boë, 2004，Role of vocal tract morphology in speech development: Perceptual targets and sensorimotor maps for synthesized French vowels from birth to adulthood. J. Speech Lang. Hear. Res. 47, 1059–1080 (2004).

L. Ménard, J.-L. Schwartz, L.-J. Boë, J. Aubin, 2007, Articulatory–acoustic relationships during vocal tract growth for French vowels: Analysis of real data and simulations with an articulatory model. J. Phon. 35, 1–19 (2007).

R. D. Kent, A. D. Murray, 1982, Acoustic features of infant vocalic utterances at 3, 6, and 9 months. J. Acoust. Soc. Am. 72, 353–365 (1982).